Conference Abstract

Not a programmer? You can mobilize data from biodiversity informatics APIs, too!

Geoffrey Ower ‡

‡ University of Illinois, Champaign, Illinois, United States of America

Abstract

An enormous amount of biodiversity information is accessible via many biodiversity informatics APIs (application programming interface). However, there may be a steep learning curve for finding, accessing, interchanging and reusing data among these APIs, often requiring a large time investment to learn a programming language, and develop and rigorously test the software. The emergence of no-code and low-code software architecture could potentially lower this barrier and empower non-programmers with the ability to use biodiversity informatics APIs. Instead of writing code, no-code/low-code applications offer flexible graphical user interfaces for quickly constructing complex data processing workflows. One such open source low-code application is n8n.io, which is a powerful workflow automation tool that could conceivably make interchanging data among scientific APIs much easier. Workflows within n8n.io are developed by connecting together nodes that take input, perform processing, and return output (Figs 1, 2). For example, you could create a workflow that starts with a node querying Wikidata for all registered external identifiers (IDs) for a species and use the output IDs to access all information about that species in biodiversity informatics APIs. After a workflow has been saved, it can be triggered manually, on a schedule, by command line, or by a remote push request. As part of the new Biology Node-to-Node (BioN2N) project, nodes were developed for several biodiversity informatics APIs including Barcode of Life, Bionomia, Catalogue of Life, ChecklistBank Datasets, Global Biodiversity Information Facility, Global Biotic Interactions, Global Names, Integrated Taxonomic Information System, Open Tree of Life, TaxonWorks, Wikidata Query Service, ZooBank, and more nodes currently are under development. A technical

demonstration will be given showing how to use n8n.io for accessing and interchanging data among biodiversity informatics APIs. A major advantage of the n8n.io platform is that it allows reusable workflows to be shared as templates, which could enable building a community of biodiversity informatics users. It also wraps hundreds of popular APIs, which could open the door to integrating scientific knowledge with many other services on the internet.
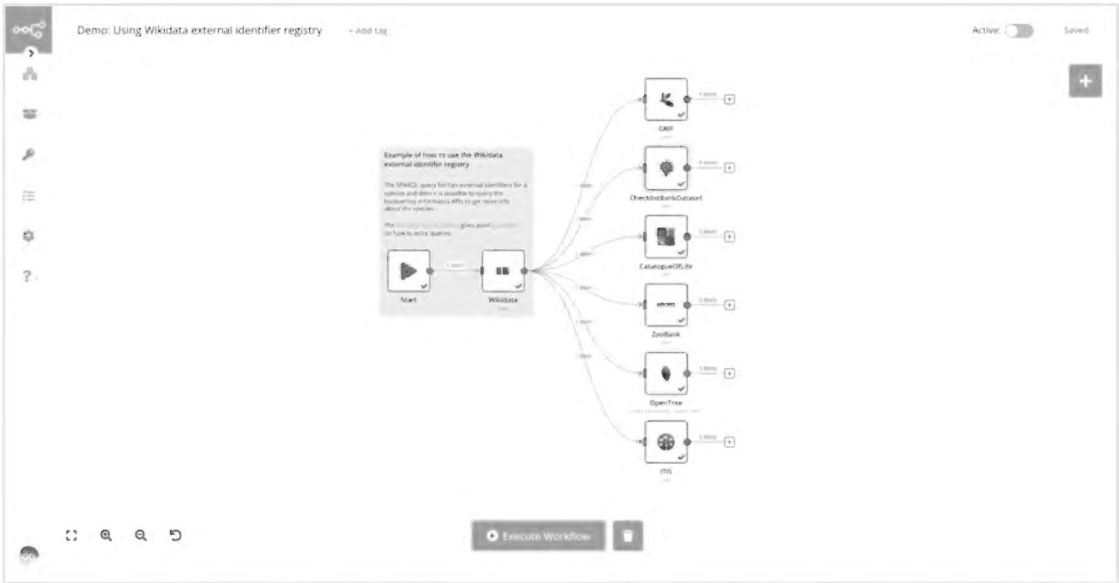

Figure 1.

This example workflow demonstrates how to query the Wikidata external identifier registry to get all biodiversity informatics identifiers for the species, *Sertularia argentea,* and in turn, query each of the APIs for data about *S. argentea.*
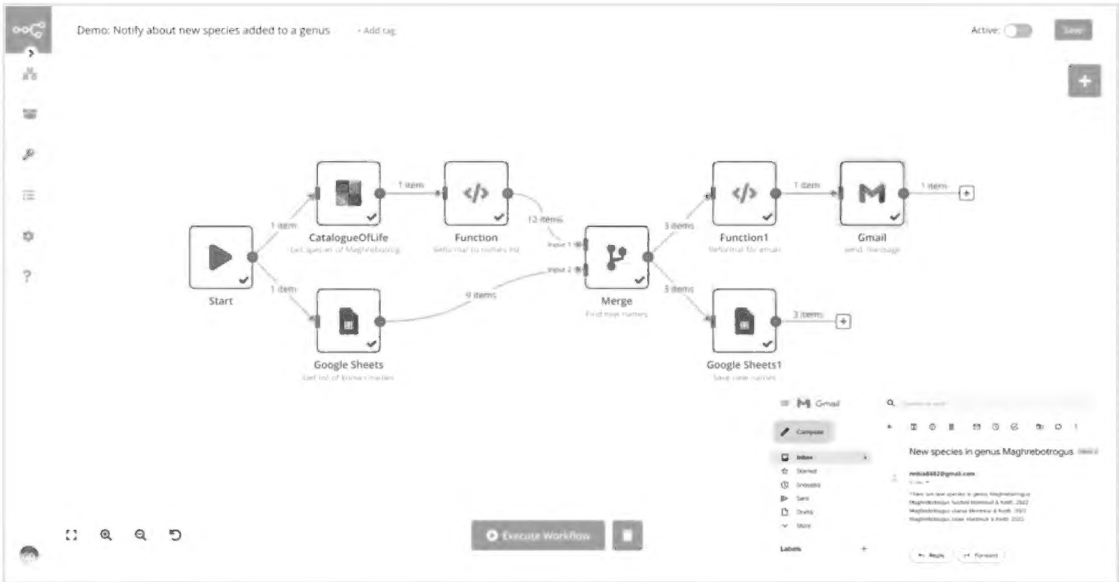

Figure 2.

In this example workflow, an email notification is sent when new species are added to the *Maghrebotrogus* genus in Catalogue of Life. The Catalogue of Life node is used to search for all species in the *Maghrebotrogus* genus and the user's list of already known species is fetched using the Google Sheets node. The data from Catalogue of Life is reformatted to match the Google Sheets data, and the Merge node is used to find any species not in the Google Sheet list of known species. When new species are added, an email notification is sent and the species are appended to the list of known species Google Sheet. The task can be run on a scheduled workflow.

Keywords

workflow, Wikidata, n8n.io, no code, low code, automation

Presenting author

Geoffrey Ower

Presented at

TDWG 2022